ICASE

FINITE ELEMENT METHODS FOR PROBLEMS
IN DYNAMIC ELASTICITY

Lois Mansfield

Report Number 77-6

Revised

May 31, 1977

(NASA-CR-185738) FINITE ELEMENT METHODS FOR PROBLEMS IN DYNAMIC ELASTICITY (ICASE) 20 p

N89-71326

Unclas

00/39 0224337

INSTITUTE FOR COMPUTER APPLICATIONS IN SCIENCE AND ENGINEERING
NASA Langley Research Center, Hampton, Virginia

Operated by the

UNIVERSITIES SPACE USRA RESEARCH ASSOCIATION

FINITE ELEMENT METHODS FOR PROBLEMS IN DYNAMIC ELASTICITY

Lois Mansfield

ABSTRACT

Finite element methods which are explicit in time are given for both the linearized and nonlinear problems in dynamic elasticity. In both cases the systems involved are written as first order hyperbolic systems. Convergence is proved for both the semi-discrete and discrete finite element methods for the linearized problem, and stability is proved for the semi-discrete finite element method for the nonlinear problem.

This report was prepared as a result of work performed under NASA Contract No. NAS1-14101 while the author was in residence at ICASE, NASA Langley Research Center, Hampton, VA. This research was also supported in part by National Science Foundation Grant MCS76-06293.

I. Introduction

In this paper we consider finite element methods to solve both the linearized and the nonlinear problems in dynamic elasticity. We are particularly interested in finite element methods with explicit time discretizations. Although the finite element method is inherently implicit, the use of explicit time discretizations in nonlinear problems produces a linear system which is independent of time to be solved at each time step. Thus, although the mass matrix must be factored, this factorization needs to be done only once. Alternatively, the linear system $Mx = b$, for M the mass matrix, can easily be solved using a few iterations of SOR, since the solution at the previous time step provides a very good initial approximation.

In the linearized problem in dynamic elasticity one seeks to find $\underline{u}$ and $\underline{\sigma}$ such that

(1) $$\frac{\partial \sigma_{ij}}{\partial x_j} + f_i = \rho \frac{\partial^2 u_i}{\partial t^2} \qquad \text{in } \Omega \quad ,$$

where

(2) $$\underline{\sigma} = E\,\underline{\varepsilon}$$

and

(3) $$\varepsilon_{ij} = \tfrac{1}{2}\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right) \quad ,$$

subject to the initial conditions

(4) $$\underline{u}(\underline{x},0) = \underline{u}_0(\underline{x}) \quad , \quad \underline{u}_t(\underline{x},0) = \underline{u}_1(\underline{x}) \quad .$$

Throughout we shall take the boundary condition to be

(5) $$\underline{u} = 0 \qquad \text{on} \quad \partial\Omega \quad .$$

Here $\underline{u}$ is the displacement vector, $\underline{\sigma}$ and $\underline{\varepsilon}$ are respectively the stress and strain tensors, E is Hooke's tensor, ρ is the density, and $\underline{f}$ is the external force. In (1) we have adopted the summation convention that the repeated subscript j in the first term indicates an implied summation over j.

To make exposition easier we shall give our results for the one dimensional scalar model problem

$$u_{tt} = u_{xx} + f \quad . \tag{6}$$

All of our results, however, have immediate extensions to (1) - (3).

The system (1) - (3) has been linearized in that the actual strain-displacement relations

$$\varepsilon_{ij} = \tfrac{1}{2}\left(\frac{\partial u_i}{\partial a_j} + \frac{\partial u_j}{\partial a_i} + \frac{\partial u_\alpha}{\partial a_i}\frac{\partial u_\alpha}{\partial a_j}\right) \tag{7}$$

have been replaced by (3) under the assumption that the displacements $u_i = x_i - a_i$ are small enough that second order terms in (7) can be neglected. For large displacements one also must distinguish between Lagrange or undeformed coordinates a_i and Euler or deformed coordinates x_i. In Lagrange coordinates, one obtains

$$\frac{\partial}{\partial a_j}\left(S_{jk}\left(\delta_{ik} + \frac{\partial u_i}{\partial a_k}\right)\right) + f_i = \rho\frac{\partial^2 u_i}{\partial t^2} \tag{8}$$

along with (2) and (7), where $\underline{S}$ is the second Piola-Kirchoff stress tensor and δ_{ik} is the Kronecker delta. Again, we shall give our results for the one dimensional scalar model equation

$$\frac{\partial}{\partial x}((1+u_x)\sigma) + f = u_{tt} \quad , \tag{9a}$$

$$\sigma = u_x + \tfrac{1}{2}(u_x)^2 \quad . \tag{9b}$$

These results have immediate extensions to (7) - (8).

The method of solution used in this paper is to write (6) and (9) as first order systems. For (6) we prove the convergence of semi-discrete and discrete Galerkin approximations. Except in special cases, e.g. linear finite elements in any dimension and smooth splines in one dimension, all on nearly uniform meshes, one obtains one order less than the optimal rate of convergence in the semi-discrete Galerkin approximation. By nearly uniform meshes, we mean meshes which are obtained by starting with an arbitrary coarse grid, or triangulation, but then always refining these intervals, or triangles, uniformly.

We prove convergence of the discrete Galerkin approximation obtained by using the staggered (half-step) variant of the leap-frog method. In addition, we show that the third order Runge-Kutta method applied to the semi-discrete system is stable, and thus convergent with third order accuracy in time (since stability + consistence = convergence). In the same way, it can be shown that the second and fourth order Runge-Kutta methods are not stable for semi-discrete Galerkin systems.

In section 3, we write the nonlinear problem (9) as a first order system. We show that the semi-discrete Galerkin approximation to this system is stable. This follows from the fact that the system conserves energy.

It has been shown in [1] that the optimal rate of convergence is obtained in semi-discrete approximations to second order hyperbolic equations. However, in elasticity it is the stresses which are of the most interest. Since the stresses are derivatives of the displacements solved for in the second order problem, the same order of convergence is obtained for the stresses in either method, except in the special cases noted above where the solution obtained by the first order system will have a higher order of convergence.

Engineers [2] have usually worked with the second order system. Since the amount of work needed to obtain the stresses at mesh points seems to be roughly the same in either method, further analytical and experimental work needs to be done to compare these two methods. Explicit finite element methods for hyperbolic equations have been considered previously by Gekeler [4] and Mock [10].

II. The Linear Problem

We let $w = u_t$, $\sigma = u_x$, and write (6), (4), and (5) as

$$
\begin{aligned}
w_t &= \sigma_x + f \\
\sigma_t &= w_x \qquad \text{in } (0,1),\ t > 0
\end{aligned}
\tag{10}
$$

subject to the initial conditions

$$
w(x,0) = u_1(x) \quad , \quad \sigma(x,0) = \frac{\partial u_0}{\partial x}(x) = \sigma_1(x) \ , \tag{11}
$$

and the boundary condition

$$
w(0,t) = w(1,t) = 0 \tag{12}
$$

For H a Banach space with norm $\|\cdot\|_H$ and $v:[0,T] \to H$ Lebesque measurable the following norms are defined

$$
\|v\|_{L^2(0,T;H)} = \left(\int_0^T \|v(\cdot,t)\|_H^2 \, dt \right)^{\frac{1}{2}}
$$

and

$$
\|v\|_{L^\infty(0,T;H)} = \sup_{0 \le t \le T} \|v(\cdot,t)\|_H \ .
$$

Let

$$
L^p(0,T;H) = \{v:[0,T] \to H:\ \|v\|_{L^p(0,T;H)} < \infty\}, \quad p = 2,\infty \ .
$$

Throughout we use $\|\cdot\|$ to denote the usual L^2 norm over $I = (0,1)$.

Choose finite dimensional subspaces $V_h \subset H_0^1(0,1)$ and $S_h \subset H^1(0,1)$. The semi-discrete Galerkin method we consider is: find $w_h \in L^2(0,T;V_h)$, $\sigma_h \in L^2(0,T;S_h)$ such that

$$(13)\quad \left(\frac{\partial w_h}{\partial t}(\cdot,t),\hat{w}_h\right) - \left(\frac{\partial \sigma_h}{\partial t}(\cdot,t),\hat{\sigma}_h\right) + \left(\sigma_h(\cdot,t)\frac{\partial \hat{w}_h}{\partial x}\right) + \left(\frac{\partial w_h}{\partial x}(\cdot,t),\hat{\sigma}_h\right)$$

$$= (f(\cdot,t),\hat{w}_h) \qquad , \qquad \text{all } (\hat{w}_h,\hat{\sigma}_h)\in V_h\times S_h, t>0.$$

$$(14a)\qquad (w_h(\cdot,0),\hat{w}_h) = (u_1,\hat{w}_h) \qquad , \qquad \text{all } \hat{w}_h \in V_h ,$$

$$(14b)\qquad (\sigma_h(\cdot,0),\hat{\sigma}_h) = (\sigma_1,\hat{\sigma}_h) \qquad , \qquad \text{all } \hat{\sigma} \in S_h .$$

We assume that the subspaces V_h and S_h have the following approximation properties. There is some $q_1 \geq 2$ such that if $v \in H^r(0,1) \cap H_0^1(0,1)$, $2 \leq r \leq q_1$, then there is some $v_h \in V_h$ such that

$$(15)\qquad \| v-v_h \|_{H^k(0,1)} \leq C\, h^r \| v \|_{H^r(0,1)}, \quad k = 0,1,$$

where C is independent of v and v_h. There is some $q_2 \geq 2$ such that if $\rho \in H^s(0,1)$, $2 \leq s \leq q_2$, then there is some $\rho_h \in S_h$ such that

$$(16)\qquad \| \rho-\rho_h \|_{H^k(0,1)} \leq C\, h^s \| \rho \|_{H^s(0,1)}, \quad k = 0,1,$$

where C is independent of ρ and ρ_h.

Theorem 1. The semi-discrete Galerkin method defined by (13) - (14) has a unique solution $w_h \in L^2(0,T;V_h)$, $\sigma_h \in L^2(0,T;S_h)$. If the exact solution (w,σ) to (10) -(12) satisfies $w \in L^\infty(0,T;H^r(0,1))$, $\partial w/\partial t \in L^2(0,T;H^r(0,1))$, and $\sigma \in L^\infty(0,T;H^s(0,1))$, $\partial\sigma/\partial t \in L^2(0,T;H^s(0,1))$, then there exists a constant $C = C(T)$ such that

$$(17)\qquad (\| w-w_h \|^2_{L^\infty(0,T;L^2(0,1))} + \| \sigma-\sigma_h \|^2_{L^\infty(0,T;L^2(0,1))})^{\frac{1}{2}} \leq Ch^\mu$$

where $\mu = \min(r-1, s-1)$ and C is independent of h.

Proof: Existence and uniqueness follows from the fact that (13) - (14) are equivalent to an intial value problem for a system of ordinary differential equations, which can easily be shown to possess a unique solution.

Let

$$\phi = w_h - \tilde{w}_h \quad , \quad \psi = \sigma_h - \tilde{\sigma}_h$$

where $w_h \in V_h$, $\tilde{\sigma}_h \in S_h$ satisfy (15) and (16) respectively.
From (13) and (10)

$$(18)\quad (\frac{\partial\phi}{\partial t}(\cdot,t),\hat{w}_h) - (\frac{\partial\psi}{\partial t}(\cdot,t),\hat{\sigma}_h) + (\psi(\cdot,t),\frac{\partial\hat{w}_h}{\partial x}) + (\frac{\partial\phi}{\partial x}(\cdot,t),\hat{\sigma}_h)$$

$$= (\frac{\partial}{\partial t}(w-\tilde{w}_h)(\cdot,t),\hat{w}_h) - (\frac{\partial}{\partial t}(\sigma-\tilde{\sigma}_h)(\cdot,t),\hat{\sigma}_h)$$

$$- (\frac{\partial}{\partial x} (\sigma-\tilde{\sigma}_h)(\cdot,t), \hat{w}_h) + (\frac{\partial}{\partial x}(w-\tilde{w}_h)(\cdot,t), \hat{\sigma}_h).$$

Choose $\hat{w}_h = \phi$, $\hat{\sigma}_h = -\psi$. This gives

$$\frac{d}{dt}\|\phi(\cdot,t)\|^2 + \frac{d}{dt}\|\psi(\cdot,t)\|^2 = 2\left(\frac{\partial}{\partial t}(w-\tilde{w}_h)(\cdot,t),\phi\right)$$

$$+ 2\left(\frac{\partial}{\partial t}(\sigma-\hat{\sigma}_h)(\cdot,t),\psi\right)$$

$$- 2\left(\frac{\partial}{\partial x}(\sigma-\hat{\sigma}_h)(\cdot,t),\phi\right)$$

$$- 2\left(\frac{\partial}{\partial x}(w-\tilde{w}_h)(\cdot,t),\psi\right) \quad .$$

Integrate from $t = 0$ to $t = \xi$ to get.

$$\|\phi(\cdot,\xi)\|^2 + \|\psi(\cdot,\xi)\|^2 \le \|\phi(\cdot,0)\|^2 + \|\psi(\cdot,0)\|^2$$

$$+ 2\sqrt{T}\,\|\phi\|_{L^\infty(0,T;L^2(0,1))}\left\|\frac{\partial(w_h-\tilde{w}_h)}{\partial t}\right\|_{L^2(0,T;L^2(0,1))}$$

$$+ 2\sqrt{T}\,\|\psi\|_{L^\infty(0,T;L^2(0,1))}\left\|\frac{\partial(\sigma-\tilde{\sigma}_h)}{\partial t}\right\|_{L^2(0,T;L^2(0,1))}$$

$$+ 2\sqrt{T}\,\|\psi\|_{L^\infty(0,T;L^2(0,1))}\left\|\frac{\partial(\sigma-\tilde{\sigma}_h)}{\partial x}\right\|_{L^2(0,T;L^2(0,1))}$$

$$+ 2\sqrt{T}\,\|\phi\|_{L^\infty(0,T;L^2(0,1))}\left\|\frac{\partial(w-\tilde{w}_h)}{\partial x}\right\|_{L^2(0,T;L^2(0,1))}$$

$$\le \|\phi(\cdot,0)\|^2 + \|\psi(\cdot,0)\|^2$$

$$+ \tfrac{1}{2}\|\phi\|^2_{L^\infty(0,T;L^2(0,1))} + 4T\left\|\frac{\partial(w-\tilde{w}_h)}{\partial t}\right\|^2_{L^2(0,T;L^2(0,1))}$$

$$+ \tfrac{1}{2}\|\psi\|^2_{L^\infty(0,T;L^2(0,1))} + 4T\left\|\frac{\partial(\sigma-\tilde{\sigma}_h)}{\partial t}\right\|^2_{L^2(0,T;L^2(0,1))}$$

$$+ 4T\left\|\frac{\partial(\sigma-\tilde{\sigma}_h)}{\partial x}\right\|^2_{L^2(0,T;L^2(\Omega))} + 4T\left\|\frac{\partial(w-\tilde{w}_h)}{\partial x}\right\|_{L^2(0,T;L^2(0,1)}$$

Now take the sup over $0 \le \xi \le T$ to obtain

$$(\|\phi\|^2_{L^\infty(0,T;L^2(0,1))} + \|\psi\|^2_{L^\infty(0,T;L^2(0,1))})^{\frac{1}{2}}$$

$$\leq \sqrt{2}\,\|\phi(\cdot,0)\| + \sqrt{2}\,\|\psi(\cdot,0)\| + 2\sqrt{T}\,\left\|\frac{\partial(w-w_h)}{\partial t}\right\|_{L^2(0,T;L^2(0,1))}$$

$$+ 2\sqrt{2T}\,\left\|\frac{\partial(\sigma-\tilde{\sigma}_h)}{\partial t}\right\|_{L^2(0,T;L^2(0,1))} + 2\sqrt{2T}\,\left\|\frac{\partial(\sigma-\tilde{\sigma}_h)}{\partial x}\right\|_{L^2(0,T;L^2(0,1))}$$

$$+ 2\sqrt{2T}\,\left\|\frac{\partial(w-\tilde{w}_h)}{\partial x}\right\|_{L^2(0,T;L^2(0,1))}$$

The inequality (17) now follows from the triangle inequality and (14).

Similarly to the results of Dupont [3] for $u_t + u_x = 0$ and the results of Lesaint [8] for positive symmetric hyperbolic systems, we obtain one less than the optimal order of convergence. In [9] it was shown that for certain subspaces, namely linear finite elements and smooth splines on nearly uniform meshes, the last two inner products in (18) are really one order of magnitude smaller than predicted by the straightforward use of the Schwartz inequality and (15) and (16). The easiest way to see this is to compare these inner products with what corresponds to the local truncation error in finite difference analysis. For details, see [9].

In the mathematical finite element literature the usual discrete time method which is analyzed is the Crank-Nicholson method. It is straightforward to show that this method when applied to (13) is second order accurate in Δt and has the accuracy in space proved in Theorem 1, independent of any ratio of Δt to h. We omit the proof. Because we are interested in using explicit methods on the nonlinear problem, we analyze a staggered step leap-frog method. Let $T = \Delta t N$ for some integer N and $v^m = v(\cdot, m\Delta t)$. This method can be described, for $0 \leq n \leq N$, as

(19a) $$(\sigma_h^{n+\frac{1}{2}}, \hat{\sigma}_h) = (\sigma_h^{n-\frac{1}{2}}, \hat{\sigma}_h) + \Delta t\left(\frac{\partial w_h^n}{\partial x}, \hat{\sigma}_h\right) \quad , \quad \text{all } \hat{\sigma}_h \in S_h$$

(19b) $$(w_h^{n+1}, \hat{w}_h) = (w_h^n, \hat{w}_h) - \Delta t\left(\sigma_h^{n+\frac{1}{2}}, \frac{\partial \hat{w}_h}{\partial x}\right) + \Delta t(f^{n+\frac{1}{2}}, \hat{w}_h) \ , \quad \text{all } \hat{w}_h \in V_h \ ,$$

where for $n = 0$, (24a) is replaced by

$$(19a') \qquad (\sigma_h^{\frac{1}{2}},\hat{\sigma}_h) = (\sigma_h^0,\hat{\sigma}_h) + \frac{\Delta t}{2}(\frac{\partial w_h^0}{\partial x},\hat{\sigma}_h) \quad , \quad \text{all } \hat{\sigma}_h \in S_h \quad ,$$

and w^0 and σ^0 are obtained from (14). On the space V_h we make the inverse assumption that

$$(20) \qquad \| \phi_x \| \le \frac{\alpha}{h} \| \phi \| \quad , \quad \text{all } \phi \in V_h$$

where the constant α is independent of ϕ and h. This assumption is satisfied by all finite element subspaces commonly in use, provided some regularity is imposed on the manner in which the mesh is refined. For example, in one dimension the ratio of the lengths of the largest to the smallest interval should remain bounded as h goes to zero. For triangulations in two dimensions there is an additional requirement that the smallest angle in any triangle stay bounded away from zero.

<u>Theorem 2.</u> Let (w,σ) be the exact solution to (10) - (12) and $\{w_h^n\}_{n=1}^N \subset V_h$, $\{\sigma^{n+\frac{1}{2}}\}_{n=0}^N \subset S_h$ be the sequences defined by (24). Suppose that $w \in L^\infty(0,T;H^r(0,1))$, $\sigma \in L^\infty(0,T;H^r(0,1))$, $\partial w/\partial t \in L^2(0,T;H^r(0,1))$, $\partial\sigma/\partial t \in L^2(0,T;H^s(0,1))$, $(\frac{\partial}{\partial t})^k w \in L^2(0,T;L^2(0,1))$, $(\frac{\partial}{\partial t})^k \sigma \in L^2(0,T;L^2(0,1))$, $k = 2,3$. Suppose that Δt is chosen so that

$$(21) \qquad \frac{\alpha\Delta t}{2h} < 1 \quad ,$$

where α is defined by (25). Then there exists a constant $C = C(T)$ such that

$$(22) \qquad (\max_{1\le n\le N} \| w^n - w_h^n \|^2 + \max_{1\le n\le N} \|\sigma^{n-\frac{1}{2}} - \sigma_h^{n-\frac{1}{2}} \|^2)^{\frac{1}{2}} \le C((\Delta t)^2 + h^\mu) \quad ,$$

where $\mu = \min(r-1,s-1)$.

Proof: Let $\phi^k = w_h^k - \tilde{w}_h^k$, $\psi^{k+\frac{1}{2}} = \sigma_h^{k+\frac{1}{2}} - \tilde{\sigma}_h^{k+\frac{1}{2}}$, where $\tilde{w}_h$ and $\tilde{\sigma}_h$ satisfy (15) and (16) respectively. From (19) and (10),

$$(\phi^{k+1}-\phi^k,\hat{w}_h) - (\psi^{k+\frac{1}{2}}-\psi^{k-\frac{1}{2}},\hat{\sigma}_h) + \Delta t(\psi^{k+\frac{1}{2}},\frac{\partial \hat{w}_h}{\partial x}) + \Delta t(\phi_x^k,\hat{\sigma}_h)$$

$$= (\Delta t w_t^{k+\frac{1}{2}} - (\tilde{w}_h^{k+1}-\tilde{w}_h^k),\hat{w}_h) - (\Delta t \sigma_t^k - (\tilde{\sigma}_h^{k+\frac{1}{2}}-\tilde{\sigma}_h^{k-\frac{1}{2}}),\hat{\sigma}_h)$$

$$- \Delta t(\frac{\partial}{\partial x}(\sigma-\tilde{\sigma}_h),\hat{w}_h) + \Delta t(\frac{\partial}{\partial x}(w-\tilde{w}_h), \hat{\sigma}_h).$$

Choose $w_h = \phi^{k+1} + \phi^k$, $\hat{\sigma}_h = -(\psi^{k+\frac{1}{2}}+\psi^{k-\frac{1}{2}})$. This gives

$$\| \phi^{k+1}\|^2 - \| \phi^k \|^2 + \| \psi^{k+\frac{1}{2}} \|^2 - \| \psi^{k-\frac{1}{2}} \|^2 + \Delta t(\psi^{k+\frac{1}{2}},\phi_x^{k+1}) - \Delta t(\psi^{k-\frac{1}{2}},\phi_x^k)$$

$$= \Delta t(\varepsilon^k,\phi^{k+1}+\phi^k) + \Delta t(\delta^k, \psi^{k+\frac{1}{2}}+\psi^{k-\frac{1}{2}}) - \Delta t(\gamma^k,\psi^{k+\frac{1}{2}}+\psi^{k-\frac{1}{2}}) - \Delta t(\nu^k,\phi^{k+1}+\phi^k)$$

where

$$\varepsilon^k = \varepsilon_1^k + \varepsilon_2^k = \left(w_t^{k+\frac{1}{2}} - \frac{w^{k+1}-w^k}{\Delta t} \right) + \left(\frac{w^{k+1}-w^k}{\Delta t} - \frac{\tilde{w}_h^{k+1}-\tilde{w}_h^k}{\Delta t} \right),$$

$$\delta^k = \delta_1^k + \delta_2^k = \left(\sigma_t^k - \frac{\sigma^{k+\frac{1}{2}}-\sigma^{k-\frac{1}{2}}}{\Delta t} \right) + \left(\frac{\sigma^{k+\frac{1}{2}}-\sigma^{k-\frac{1}{2}}}{\Delta t} - \frac{\tilde{\sigma}_h^{k+\frac{1}{2}}-\tilde{\sigma}_h^{k-\frac{1}{2}}}{\Delta t} \right),$$

$$\gamma^k = \frac{\partial}{\partial x}(w^k-\tilde{w}_h^k),$$

$$\nu^k = \frac{\partial}{\partial x}(\sigma^k-\tilde{\sigma}_h^k).$$

Summing from $k = 1$ to $k = n$ and using the Schwartz inequality along with the inequality $ab \le (\varepsilon a^2/2) + (b^2/2\varepsilon)$ for ε positive gives

$$(23) \quad \| \phi^{n+1}\|^2 + \| \psi^{n+\frac{1}{2}}\|^2 + \Delta t(\psi^{n+\frac{1}{2}},\phi_x^{n+1})$$

$$\le \| \phi^1 \|^2 + \| \psi^{\frac{1}{2}}\|^2 + \Delta t(\psi^{\frac{1}{2}},\phi_x^1) + \frac{2T}{\beta} \Delta t \sum_{k=0}^{n} \| \varepsilon^k \|^2$$

$$+ \frac{2T}{\beta} \Delta t \sum_{k=1}^{n} \| \delta^k \|^2 + \frac{2T}{\beta} \Delta t \sum_{k=1}^{n} \| \gamma^k \|^2 + \frac{2T}{\beta} \Delta t \sum_{k=1}^{n} \| \nu^k \|^2$$

$$+ \beta \frac{\Delta t}{2T} \sum_{k=1}^{n} \| \phi^{k+1}+\phi^k \|^2 + \beta \frac{\Delta t}{2T} \sum_{k=1}^{n} \| \psi^{k+\frac{1}{2}}+\psi^{k-\frac{1}{2}} \|^2 .$$

where $0 < \beta < 1$. Using (24b) and (24a') we obtain

$$(24) \qquad \| \phi^1 \|^2 + 2\| \psi^{\frac{1}{2}} \|^2 = \| \phi^0 \|^2 + 2\| \psi^0 \|^2 - \Delta t(\psi^{\frac{1}{2}},\phi^1_x) + \Delta t(\psi^0,\phi^0_x)$$

$$+ \Delta t(\varepsilon^0,\phi^1+\phi^0) + \Delta t(\hat{\delta}^0,\psi^{\frac{1}{2}}+\psi^0) - \Delta t(\gamma^0,\psi^{\frac{1}{2}}+\psi^0) - \Delta t(\nu^0,\phi^1+\phi^0)$$

where

$$\hat{\delta}^0 = \hat{\delta}^0_1 + \hat{\delta}^0_2 = (\sigma^0_t - 2(\frac{\sigma^{\frac{1}{2}}-\sigma^0}{\Delta t})) + \frac{2}{\Delta t}(\sigma^{\frac{1}{2}}-\tilde{\sigma}^{\frac{1}{2}}_h+\tilde{\sigma}^0_h-\sigma^0_h) .$$

Then

$$(25a) \qquad \| \delta^k_1 \|^2 = \int_0^1 [\int_{t^{k-\frac{1}{2}}}^{t^{k+\frac{1}{2}}} \{(t^k-\tau)^2_+ - (\frac{t^{n+\frac{1}{2}}-\tau}{\Delta t})^2\} \frac{\partial^3\sigma}{\partial t^3}(\cdot,\tau)d\tau]^2$$

$$\leq K(\Delta t)^3 \int_{t^{k-\frac{1}{2}}}^{t^{k+\frac{1}{2}}} \| \frac{\partial^3\sigma}{\partial t^3}(\cdot,\tau) \|^2 ,$$

$$(25b) \qquad \| \delta^k_2 \|^2 = \frac{1}{\Delta t} \int_{t^{k-\frac{1}{2}}}^{t^{k+\frac{1}{2}}} \frac{\partial}{\partial t}(\sigma-\tilde{\sigma}_h)(\cdot,\tau)d\tau \leq \frac{1}{\Delta t} \int_{t^{k-\frac{1}{2}}}^{t^{k+\frac{1}{2}}} \| \frac{\partial}{\partial t}(\sigma-\tilde{\sigma}_h)(\cdot,\tau) \|^2 ,$$

$$(25c) \qquad \| \hat{\delta}^0_1 \| \leq K(\Delta t)^2 \int_{t^0}^{t^{\frac{1}{2}}} \| \frac{\partial^2\sigma}{\partial t}(\cdot,\tau) \|^2 ,$$

where K is some constant which is independent of Δt and σ. Similar estimates hold for $\varepsilon^k_1,\varepsilon^k_2$, and δ^0_2.

Substituting (24) and (25) into (23), and using (20) gives

$$(1-\frac{\alpha\Delta t}{2h}) \left[\| \phi^{n+1} \|^2 + \| \psi^{n+\frac{1}{2}} \|^2\right] \leq (1+\beta+\frac{\alpha\Delta t}{2h}) \left[\| \phi^0 \|^2 + \| \psi^0 \|^2\right]$$

$$+ K\Bigg[(\Delta t)^4 \Bigg\{\| \frac{\partial^3 w}{\partial t^3} \|^2_{L^2(0,T;L^2(0,1))} + \| \frac{\partial^3\sigma}{\partial t^3} \|^2_{L^2(0,T;L^2(0,1))} + \| \frac{\partial^2\sigma}{\partial t^2} \|^2_{L^2(0,T;L^2(0,1))}$$

$$+ \| \frac{\partial}{\partial t}(\sigma-\tilde{\sigma}_h) \|^2_{L^2(0,T;L^2(0,1))} + \| \frac{\partial}{\partial t}(w-\tilde{w}_h) \|^2_{L^2(0,T;L^2(0,1))} + \max_{0\leq k\leq N} \| \frac{\partial}{\partial x}(w-\tilde{w}_n) \|^2$$

$$+ \max_{0\leq k\leq N} \| \frac{\partial}{\partial x} (\sigma^k-\tilde{\sigma}^k_h) \|^2\Bigg] + \beta \max_{1\leq k\leq N} \| \phi^k \|^2 + \beta \max_{0\leq k<N-1} \| \psi^{k+\frac{1}{2}} \|^2 .$$

Taking the sup over k and using the triangle inequality along with (15) and (16) gives (22) and proves the theorem.

The optimal order of convergence in h can be obtained for certain subspaces in the same way as for the semi-discrete approximation. We now show that the stability requirement (20) is the same as that given in Bathe and Wilson [2, Chapter 9]. For the general dynamic elasticity problem, the system (10) becomes

$$\begin{aligned} w_t &= T^*\underline{\sigma} + \underline{f} \\ \underline{\sigma}_t &= T\underline{w} \end{aligned} \tag{26}$$

For (26), the inequality (20) becomes

$$\| T\phi \| \leq \frac{\alpha}{h} \| \phi \| \quad , \qquad \text{all } \phi \varepsilon V_h \quad ,$$

or

$$\frac{\alpha}{h} = \sup_{\phi \varepsilon V_h} \frac{\| T\phi \|}{\| \phi \|} .$$

Let $\{\phi_i\}_{i=1}^n$ be a basis for V_h. Then for $\phi_h \varepsilon V_h$,

$$\phi_h = \sum_{i=1}^n a_i \phi_i$$

so that

$$\frac{\| T\phi \|^2}{\| \phi \|^2} = \frac{a^T K a}{a^T M a} \leq \lambda_{max}$$

and thus

$$\frac{\alpha}{h} = \sqrt{\lambda_{max}} \tag{27}$$

Here K and M are respectively the "stiffness" matrix and the "mass" matrix, and λ_{max} is the largest eigenvalue of $\lambda Mv = Kv$. This agrees with what Bathe and Wilson, [2, page 353], determined for the centered difference method applied to the second order equation.

We now show that the third order Runge-Kutta method when applied to the semi-discrete Galerkin approximation is stable. Convergence can then be shown by applying the classical theorem that stability plus consistency equals convergence. The same techniques can be used to show that the second and fourth order Runge-Kutta methods applied to the semi-discrete Galerkin approximation are unstable. The third order Runge-Kutta method can be written as the three step procedure (see [6, page 43])

$$
\begin{aligned}
u^{n+\frac{1}{3}} &= u^n + \frac{\Delta t}{3}(u')^n \quad , \\
u^{n+\frac{2}{3}} &= u^n + \frac{2\Delta t}{3}(u')^{n+\frac{1}{3}} \quad , \qquad (28) \\
u^{n+1} &= u^n + \frac{\Delta t}{4}(3(u')^{n+\frac{2}{3}} + (u')^n) \ .
\end{aligned}
$$

Let $\underline{u}^n$ be the vector $[\underline{\alpha}^n,\underline{\beta}^n]^T$, where α_i^n and β_j^n are the coefficients in the expansions

$$
w^n = \sum_{i=1}^{m_1} \alpha_i^n \hat{\phi}_i \quad , \quad \sigma^n = \sum_{j=1}^{m_2} \beta_j^n \hat{\psi}_j
$$

where $\{\hat{\phi}_i\}_{i=1}^{m_1}$, and $\{\hat{\psi}_i\}_{i=1}^{m_2}$ are orthonormal bases with respect to the inner product $(\cdot,\cdot)$ for V_h and S_h respectively. For the system (13), (28) can be expressed (for $f = 0$) as

$$
\underline{u}^{n+1} = \underline{u}^n + \Delta t B\underline{u}^n + \frac{(\Delta t)^2}{2} B^2\underline{u}^n + \frac{(\Delta t)^3}{6} B^3\underline{u}^n \qquad (29)
$$

where

$$
B = \begin{pmatrix} 0 & C \\ -C^T & 0 \end{pmatrix}
$$

where C is the rectangular matrix with elements $c_{ij} = ((\hat{\psi}_i)_x, \hat{\phi}_j)$.

Note that

$$\| w^n \|^2 + \| \sigma^n \|^2 = (\underline{u}^n, \underline{u}^n) = |||\underline{u}^n|||^2 .$$

Theorem 3. Let $\{\underline{u}^k\}$ be the sequence defined by (29) where $\underline{u}^0$ is obtained from (14). Then

$$|||\underline{u}^{n+1}||| \leq |||\underline{u}^0|||^2$$

provided that

$$\frac{\Delta t}{\sqrt{3}} |||B||| < 1 \tag{30}$$

Proof. Take the $|||\cdot|||$ norm of both sides of (29). Because B is skew-symmetric, most of the terms on the right side drop out and we are left with

$$\begin{aligned} |||\underline{u}^{n+1}|||^2 &= |||\underline{u}^n|||^2 - \frac{(\Delta t)^4}{12} |||B^2\underline{u}^n|||^2 + \frac{(\Delta t)^6}{36} |||B^3\underline{u}^n|||^2 \\ &\leq |||\underline{u}^n|||^2 + \frac{(\Delta t)^4}{12}\left(\frac{(\Delta t)^2}{3} |||B|||^2 - 1\right) |||B^2\underline{u}^n|||^2 \\ &\leq |||\underline{u}^n|||^2 \end{aligned}$$

if (30) holds. The theorem now follows by recursion.

III. The Nonlinear Problem

We let $w = u_t$, $\sigma = u_x + \frac{1}{2}(u_x)^2$ and write (9), (4), and (5) as

$$\begin{aligned} w_t &= \frac{\partial}{\partial x}((1+u_x)\sigma) + f , \\ \sigma_t &= (1+u_x)w_x \quad \text{in} \quad (0,1) , \quad t>0 , \end{aligned} \tag{31}$$

subject to the initial conditions

$$(32)\qquad w(x,0) = u_1(x) \quad , \quad \sigma(x,0) = \frac{\partial u_0}{\partial x}(x) + \frac{1}{2}\left(\frac{\partial u_0}{\partial x}\right)^2 = \sigma_1(x) \quad ,$$

and the boundary condition

$$(33)\qquad w(0,t) = (w(1,t) = 0 \quad .$$

To obtain the second equation in (31), we needed to assume that we could interchange time and space differentiations.

The system (31) - (33) can be written in weak form as

$$(34)\qquad (w_t(\cdot,t),\hat{w}) - (\sigma_t(\cdot,t),\hat{\sigma}) + ((1+u_x(\cdot,t))\sigma(\cdot,t),\hat{w}_x)$$

$$+ ((1+u_x(\cdot,t))w_x(\cdot,t),\hat{\sigma}) = (f(\cdot,t),\hat{w}) \quad ,$$

$$(35a)\qquad (w(\cdot,0),\hat{w}) = (u_1,\hat{w}) \quad ,$$

$$(35b)\qquad (\sigma(\cdot,0),\hat{\sigma}) = (\sigma_1,\hat{\sigma}) \quad ,$$

$$\text{all} \quad (\hat{w},\hat{\sigma}) \in H_0^1(0,1) \times H^1(0,1).$$

The semi-discrete Galerkin method we consider is: find $w_h \in L^2(0,T;V_h)$, $\sigma_h \in L^2(0,T;S_h)$ such that

$$(36)\qquad (\frac{\partial w_h}{\partial t}(\cdot,t),\hat{w}_h) - (\frac{\partial \sigma_h}{\partial t}(\cdot,t),\hat{\sigma}_h) + ((1+u_x(\cdot,t))\sigma_h(\cdot,t),\frac{\partial \hat{w}_h}{\partial x})$$

$$+ ((1+u_x(\cdot,t))\frac{\partial w_h}{\partial x}(\cdot,t),\hat{\sigma}_h) = (f(\cdot,t),\hat{w}_h) \quad ,$$

$$(37a)\qquad (w_h(\cdot,0),\hat{w}_h) = (u_1,\hat{w}_h) \quad ,$$

$$(37b)\qquad (\sigma_h(\cdot,0),\hat{\sigma}_h) = (\sigma_1,\hat{\sigma}_h) \quad ,$$

$$\text{all} \quad (w_h,\sigma_h) \in V_h \times S_h \quad .$$

<u>Theorem 4.</u> Let (w,σ) be the solution to (34) - (35) and (w_h,σ_h) be the semi-discrete Galerkin solution to (36) - (37). At any time ξ such that $0 \leq \xi \leq T$

$$\|w(\cdot,\xi)\|^2 + \|\sigma(\cdot,\xi)\|^2 \leq 2(\|w(\cdot,\xi)\|^2 + \|\sigma(\cdot,0)\|^2) + 4T\|f\|^2_{L^2(0,T;L^2(0,1))}, \tag{38}$$

$$\|w_h(\cdot,\xi)\|^2 + \|\sigma_h(\cdot,\xi)\|^2 \leq 2(\|w_h(\cdot,0)\|^2 + \|\sigma_h(\cdot,0)\|^2) + 4T\|f\|^2_{L^2(0,T;L^2(0,1))}, \tag{39}$$

holds for the forced vibration problem. For free vibration problems (f=0), one has

$$\|w(\cdot,\xi)\|^2 + \|\sigma(\cdot,\xi)\|^2 = \|w(\cdot,0)\|^2 + \|\sigma(\cdot,0)\|^2, \tag{40}$$

$$\|w_h(\cdot,\xi)\|^2 + \|\sigma_h(\cdot,\xi)\|^2 = \|w_h(\cdot,0)\|^2 + \|\sigma_h(\cdot,0)\|^2, \quad 0 \leq \xi \leq T. \tag{41}$$

<u>Proof.</u> Let $\hat{w} = w$, $\hat{\sigma} = -\sigma$ in (34). This gives

$$\frac{d}{dt}\|w(\cdot,t)\|^2 + \frac{d}{dt}\|\sigma(\cdot,t)\|^2 = 2(f(\cdot,t),\hat{w}).$$

Integrate from $t = 0$ to $t = \xi$ to obtain

$$\| w(\cdot,\xi) \|^2 + \| \sigma(\cdot,\xi) \|^2 - \| w(\cdot,0) \|^2 - \| \sigma(\cdot,0) \|^2$$

$$= 2 \int_0^\xi (f(\cdot,t),w)$$

$$\le 2 \sqrt{T} \; \|w\|_{L^\infty(0,T;L^2(0,1))} \; \| f \|_{L^2(0,T;L^2(0,1))}$$

$$\le \tfrac{1}{2} \| w\|^2_{L^\infty(0,T;L^2(0,1))} + 2T \| f\|^2_{L^2(0,T;L^2(0,1))} .$$

Taking the sup over all ξ such that $0 \le \xi \le T$ gives (38). The inequality (39) and the identities (40) and (41) are derived in the same way.

Note that the identities (40) and (41) indicate that both the weak system (34) - (35) and the semi-discrete Galerkin system (36) = (37) conserve energy (for $f = 0$). This was true also of the linear problem.

It has been observed [11, page 128] and [5] that the use of leap frog in nonlinear problems can cause difficulties because of nonlinear instability. For a system like (31) where the semi-discrete approximation is conservative this nonlinear instability seems to occur only when the coefficient $1 + u_x$ oscillates around zero. See [5, page 207]. One can avoid this instability by adding a small dissipative term to the left hand side. Thus, one replaces the mass matrix M by $M + \alpha(\Delta t)^2 K$, where K is the stiffness matrix, and α is some constant. If this is done only to (19b) the procedure can be thought of as adding artificial damping. It can be shown that this does not reduce the order of convergence. The addition of artificial damping can also be used as a way of treating shocks analogously to artificial viscosity for inviscid fluid flows. An excellent discussion of the motivation behind artificial viscosity is given in Lax [7].

References

[1] Baker, G. A., "Error estimates for finite element methods for second order hyperbolic equations," SIAM Journal on Numerical Analysis, Vol. 13 (1976), 564-76.

[2] Bathe, K.-J. and E. L. Wilson, NUMERICAL METHODS IN FINITE ELEMENT ANALYSIS, Prentice-Hall, Englewood Cliffs, NJ, 1976.

[3] Dupont, T., "Galerkin methods for first order hyperbolics: an example," SIAM Journal on Numerical Analysis, Vol. 10 (1973), 898-907.

[4] Gekeler, E., "Galerkin-Runge-Kutta methods and hyperbolic initial boundary value problems," Computing, 18 (1977), 79-88.

[5] Kreiss, H.-O. and J. Oliger, "Comparison of accurate methods for the integration of hyperbolic equations," Tellus, Vol. 24 (1972), 199-215.

[6] Lapidus, L. and J. H. Steinfeld, NUMERICAL SOLUTION OF ORDINARY DIFFERENTIAL EQUATIONS, Academic Press, New York, 1971.

[7] Lax, P. D., "Nonlinear partial differential equations and computing," SIAM Review, Vol.11 (1969), 7-19.

[8] Lesaint, P., "Finite element methods for symmetric hyperbolic equations," Numer. Math., Vol. 21 (1973), 244-55.

[9] Mansfield, L., "Mixed finite element methods for elliptic equations," ICASE Report No. 76-24, 1976.

[10] Mock, M. S., "Explicit finite element schemes for first order symmetric hyperbolic systems," Numer. Math., Vol. 26 (1976), 367-78.

[11] Richtmyer, R. D. and K. W. Morton, DIFFERENCE METHODS FOR INITIAL-VALUE PROBLEMS, Second Edition, Interscience, New York, 1967.